d his

(FILE 'HOME' ENTERED AT 11:21:35 ON 07 SEP 2004)

FILE 'CAPLUS' ENTERED AT 11:25:08 ON 07 SEP 2004
L1 450 S (PROTEIN (5W)((X(2W)RAY) OR CRYSTAL?)) AND (MOLECULAR REPLACE
L2 71 S L1 AND (SEARCH MODEL)

=> d bib, abs 19,24,25

L2 ANSWER 19 OF 71 CAPLUS COPYRIGHT 2004 ACS on STN
AN 2001:695651 CAPLUS
DN 135:368799
TI How to take advantage of non-crystallographic symmetry in **molecular replacement**: 'locked' rotation and translation functions
AU Tong, Liang
CS Department of Biological Sciences, Columbia University, New York, NY, 10027, USA
SO Acta Crystallographica, Section D: Biological Crystallography (2001), D57(10), 1383-1389
CODEN: ABCRE6; ISSN: 0907-4449
PB Munksgaard International Publishers Ltd.
DT Journal
LA English
AB Many protein mols. form assemblies that obey point-group symmetry. These assemblies are often situated at general positions in the unit cell such that the point-group symmetry of the assembly becomes non-crystallog. symmetry (NCS) in the crystal. The presence of NCS places significant constraints on structure determination by the **mol.-replacement** method. The locked rotation and translation functions have been developed to take advantage of the presence of NCS in this structure determination, which generally requires four steps. (i) The locked self-rotation function is used to determine the orientation of the NCS assembly in the crystal, relative to a pre-defined 'standard' orientation of this NCS point group. (ii) The locked cross-rotation function is used to determine the orientation of one monomer of the assembly in the standard orientation. This calcn. requires only the structure of the monomer as the **search model**. (iii) The locked translation function is used to determine the position of this monomer relative to the center of the assembly. Information obtained from steps (ii) and (iii) will produce a model of the entire assembly centered at the origin of the coordinate system. (iv) An ordinary translation function is used to determine the center of the assembly in the crystal unit cell, using as the **search model** the structure of the entire assembly produced in step (iii). The locked rotation and translation functions simplify the structure-determination process in the presence of NCS. Instead of searching for each monomer sep., the locked calcns. search for a single rotation or translation. Moreover, the locked functions reduce the noise level in the calcn., owing to the averaging over the NCS elements, and increase the signals as all monomers of the assembly are taken into account at the same time.
RE.CNT 13 THERE ARE 13 CITED REFERENCES AVAILABLE FOR THIS RECORD
ALL CITATIONS AVAILABLE IN THE RE FORMAT

L2 ANSWER 24 OF 71 CAPLUS COPYRIGHT 2004 ACS on STN
AN 2000:847724 CAPLUS
DN 134:128180
TI An approach to multi-copy search in **molecular replacement**
AU Vagin, Alexei; Teplyakov, Alexei
CS Department of Chemistry, University of York, Heslington, York, YO1 5DD, UK
SO Acta Crystallographica, Section D: Biological Crystallography (2000), D56(12), 1622-1624

CODEN: ABCRE6; ISSN: 0907-4449
PB Munksgaard International Publishers Ltd.
DT Journal
LA English
AB The **mol.-replacement** method has been extended to a simultaneous search for multiple copies of the macromol. in the unit cell. The central point of this approach is the construction of a multi-copy **search model** from the properly oriented monomers using a special translation function. The multi-copy search method has been implemented in the program MOLREP and successfully tested using exptl. data.
RE.CNT 15 THERE ARE 15 CITED REFERENCES AVAILABLE FOR THIS RECORD
ALL CITATIONS AVAILABLE IN THE RE FORMAT

L2 ANSWER 25 OF 71 CAPLUS COPYRIGHT 2004 ACS on STN
AN 2000:807192 CAPLUS
DN 134:112341
TI Does NMR Mean "Not for **Molecular Replacement**"? Using NMR-Based **Search Models** to Solve **Protein Crystal** Structures
AU Chen, Y. W.; Dodson, E. J.; Kleywegt, G. J.
CS Centre for Protein Engineering and Cambridge University Chemical Laboratory, MRC Centre, Cambridge, CB2 2QH, UK
SO Structure (London) (2000), 8(11), R213-R220
CODEN: STRUE6; ISSN: 0969-2126
PB Elsevier Science Ltd.
DT Journal; General Review
LA English
AB A review with 47 refs. The test cases discussed in this study show that using NMR models to search for MR solns. is now quite feasible, at least in favorable circumstances. Modern NMR studies now provide models which are more similar to those found by crystallog. techniques, indicating that the protein folds found in the solution usually closely resemble those in the crystal and helping to scotch the belief that the crystal environment distorts the protein. Techniques developed for utilizing NMR models should be valid for performing MR studies with distantly homologous proteins. This could prove to be a valuable tool for structural genomics. However, MR techniques still do not guarantee success and further studies are required to fully exploit this method.
RE.CNT 47 THERE ARE 47 CITED REFERENCES AVAILABLE FOR THIS RECORD
ALL CITATIONS AVAILABLE IN THE RE FORMAT

FILE 'WPIDS' ENTERED AT 11:43:02 ON 07 SEP 2004
L3 22 S (PROTEIN (5W)((X(2W)RAY) OR CRYSTAL?)) AND (MOLECULAR REPLACE

=> d bib,kwic 11-22

L3 ANSWER 11 OF 22 WPIDS COPYRIGHT 2004 THOMSON DERWENT on STN
AN 2003-403195 [38] WPIDS
DNN N2003-321578 DNC C2003-107403
TI New S8 protein defined from Staphylococcus aureus, useful for identifying inhibitors of the rRNA-binding activity of S. aureus S8, and in screening of molecules and/or designing of new molecules that bind to the S8 protein structure.
DC B04 D16 S03 T01
IN CONCHA, N O; GONTAREK, R R; JANSON, C A
PA (SMIK) SMITHKLINE BEECHAM CORP
CYC 101
PI WO 2003033531 A1 20030424 (200338)* EN 41
RW: AT BE BG CH CY CZ DE DK EA EE ES FI FR GB GH GM GR IE IT KE LS LU MC MW MZ NL OA PT SD SE SK SL SZ TR TZ UG ZM ZW
W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ OM PH PL PT RO RU SD SE SG SI SK SL TJ TM TN TR TT TZ UA UG US UZ VC VN YU ZA ZM ZW
ADT WO 2003033531 A1 WO 2002-US32859 20021015
PRAI US 2001-329439P 20011015
AB . . .
a protein having the coordinates listed in the specification;
(2) a heavy atom derivative of a S. aureus S8 **protein crystal**, where the rRNA-binding function comprises a protein having the coordinates listed in the specification;
(3) a process for identifying an. . . crystal or its portions, to determine a crystal form of a mutant, homolog or co-complex of the rRNA-binding function by **molecular replacement**;
(6) a process for designing drugs for inhibiting S. aureus S8 activity using the atomic coordinates of a S. aureus. . .
TECH. . .
of S8 lined by residues 4-6, 30-32, 56-57, 82-92, 107-111, and 122-125 that interact with nucleotides A587-A758. The S8 rRNA-binding **protein** in **crystalline** form has lattice constants of a = 42.1Angstrom, b = 55.9Angstrom, c = 61.3Angstrom, alpha = 09.0degrees, beta = 09.0degrees,. . .

L3 ANSWER 12 OF 22 WPIDS COPYRIGHT 2004 THOMSON DERWENT on STN
AN 2003-247867 [24] WPIDS
DNC C2003-063733
TI Novel 2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase protein or its functional **protein** subunit, in **crystalline** form, useful for identifying and designing inhibitors and activators of the protein.
DC B04 C06 D16
IN BUCHANAN, S G; GAJIWALA, K S; LOUIE, G V; SAUDER, J M; SAUDER, M J
PA (STRU-N) STRUCTURAL GENOMIX; (BUCH-I) BUCHANAN S G; (GAJI-I) GAJIWALA K S; (LOUI-I) LOUIE G V; (SAUD-I) SAUDER J M; (STRU-N) STRUCTURAL GENOMIX INC
CYC 100
PI WO 2002102991 A2 20021227 (200324)* EN 370
RW: AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW MZ NL OA PT SD SE SL SZ TR TZ UG ZM ZW
W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ OM PH PL PT RO RU SD SE SG SI SK SL TJ TM TN TR TT TZ UA UG US UZ VN YU ZA ZM ZW

US 2003073134 A1 20030417 (200329)
AU 2002322265 A1 20030102 (200452)

ADT WO 2002102991 A2 WO 2002-US19451 20020617; US 2003073134 A1 Provisional US 2001-299058P 20010618, US 2002-174410 20020617; AU 2002322265 A1 AU 2002-322265 20020617

FDT AU 2002322265 A1 Based on WO 2002102991

PRAI US 2001-299058P 20010618; US 2002-174410 20020617

TI Novel 2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase protein or its functional **protein** subunit, in **crystalline** form, useful for identifying and designing inhibitors and activators of the protein.

AB WO2002102991 UPAB: 20030410

NOVELTY - A 2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (MECPS) protein (I) or a functional MECPS **protein** subunit, in **crystalline** form, is new.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for the following:

(1) Producing (M1) a. . .

TECH. . .

binding pocket of a MECPS protein, by obtaining 3D structural coordinates defining the protein or a binding pocket of the **protein**, from a **crystal** of the protein; and

(b) introducing the structural coordinate into a computer to produce a database containing the molecular structural coordinates of the protein or binding pocket.

M2 comprises:

(a) generating a representation of binding pocket of a MECPS **protein** in a co-**crystal** with a compound, preferably a compound rationally designed to be capable of binding the binding pocket by preparing a binding. . . MECPS active site or binding pocket; and

(c) determining whether the potential modulator activates or inhibits the activity of the **protein**.

M5 comprises:

(a) generating an **X-ray** diffraction pattern from a crystallized form of the molecule or molecular complex, using a **molecular replacement** method to interpret the structure of the molecule, where the **molecular replacement** method uses the structural coordinates given in the specification, or its subset comprising a binding pocket, where the structural coordinates. .

.

TT TT: NOVEL METHYL ERYTHRITOL SYNTHASE **PROTEIN** FUNCTION **PROTEIN CRYSTAL** FORM USEFUL IDENTIFY DESIGN INHIBIT ACTIVATE PROTEIN.

L3 ANSWER 13 OF 22 WPIDS COPYRIGHT 2004 THOMSON DERWENT on STN

AN 2003-247844 [24] WPIDS

DNN N2003-197048 DNC C2003-063714

TI New pyrazolo(3,4-c)pyridazine derivatives are glucogen synthase kinase -3 inhibitors useful for treating e.g. schizophrenia, Alzheimer's disease, diabetes, autoimmune diseases, allergy, asthma, multiple sclerosis, and baldness.

DC B02 B04 S03 T01

IN ARNOST, M J; GREEN, J; HAAR, E T; SWENSON, L; TER HAAR, E

PA (ARNO-I) ARNOST M J; (GREE-I) GREEN J; (HAAR-I) HAAR E T; (SWEN-I) SWENSON L; (VERT-N) VERTEX PHARM INC

CYC 101

PI WO 2002088078 A2 20021107 (200324)* EN 778

RW: AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW MZ NL OA PT SD SE SL SZ TR TZ UG ZM ZW

W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ OM PH PL PT RO RU SD SE SG SI SK SL TJ TM TN TR TT TZ UA UG US UZ VN YU ZA ZM ZW

US 2003125332 A1 20030703 (200345)
AU 2002259071 A1 20021111 (200433)
EP 1435957 A2 20040714 (200446) EN
R: AL AT BE CH CY DE DK ES FI FR GB GR IE IT LI LT LU LV MC MK NL PT RO SE SI TR

ADT WO 2002088078 A2 WO 2002-US13511 20020429; US 2003125332 A1 Provisional US 2001-287366P 20010430, Provisional US 2001-297094P 20010608, Provisional US 2002-361899P 20020227, US 2002-135255 20020429; AU 2002259071 A1 AU 2002-259071 20020429; EP 1435957 A2 EP 2002-729056 20020429, WO 2002-US13511 20020429

FDT AU 2002259071 A1 Based on WO 2002088078; EP 1435957 A2 Based on WO 2002088078

PRAI US 2002-361899P 20020227; US 2001-287366P 20010430;
US 2001-297094P 20010608; US 2002-135255 20020429

AB . . .
complex comprising (C2) involves:
(i) producing and purifying GSK-3 beta protein;
(ii) mixing a crystallization solution with the **protein** complex to produce a **crystallizable** composition; and
(iii) crystallizing the composition;
(4) A molecule or molecular complex comprises a binding pocket defined by. . .

TECH. . .
a display terminal, a printer or disk drive.

TECHNOLOGY FOCUS - BIOLOGY - Preferred Components: (C2) is HSSPHQpSEDEEE. The GSK-3beta **protein** in the **crystal** is selected from 420 amino acid sequences as given in the specification, amino acid residues 7 - 420 of the. . .

L3 ANSWER 14 OF 22 WPIDS COPYRIGHT 2004 THOMSON DERWENT on STN
AN 2003-229481 [22] WPIDS
CR 2003-221757 [21]
DNN N2003-182547 DNC C2003-059031
TI Novel perosamine synthase homolog protein or its functional **protein** subunit, in a **crystalline** form, useful for identifying and designing inhibitors and activators of the protein, and for designing antimicrobials.
DC B04 D16 T01
IN BADGER, J; BUCHANAN, S G; HANS-JOACHIM, M; HENDLE, J; NOLAND, B; MULLER-DIECKMANN, H
PA (BADG-I) BADGER J; (BUCH-I) BUCHANAN S G; (HEND-I) HENDLE J; (MULL-I) MULLER-DIECKMANN H; (NOLA-I) NOLAND B; (STRU-N) STRUCTURAL GENOMIX INC
CYC 100
PI WO 2003006617 A2 20030123 (200322)* EN 424
RW: AT BE BG CH CY CZ DE DK EA EE ES FI FR GB GH GM GR IE IT KE LS LU MC MW MZ NL OA PT SD SE SK SL SZ TR TZ UG ZM ZW
W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ OM PH PL PT RO RU SD SE SG SI SK SL TJ TM TN TR TT TZ UA UG US UZ VN YU ZA ZM ZW
US 2003101005 A1 20030529 (200337)
AU 2002332410 A1 20030129 (200452)

ADT WO 2003006617 A2 WO 2002-US21935 20020712; US 2003101005 A1 Provisional US 2001-305428P 20010713, US 2002-194728 20020712; AU 2002332410 A1 AU 2002-332410 20020712

FDT AU 2002332410 A1 Based on WO 2003006617

PRAI US 2001-305428P 20010713; US 2002-194728 20020712

TI Novel perosamine synthase homolog protein or its functional **protein** subunit, in a **crystalline** form, useful for identifying and designing inhibitors and activators of the protein, and for designing antimicrobials.

AB WO2003006617 UPAB: 20040813

NOVELTY - An perosamine synthase homolog (PSH) protein (I), or a functional subunit of PSH **protein**, in its **crystalline** form, is new.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for the following:

(1) producing (M1) a. . . pocket of a PSH protein, by:

(a) obtaining 3D structural coordinates defining the protein or a binding pocket of the **protein**, from a **crystal** of the protein; and

(b) introducing the structural coordinate into a computer to produce a database containing the molecular. . . produced by M1;

(3) producing (M2) a computer readable database comprising a representation of a binding pocket of a PSH **protein** in a co-**crystal** with a compound, optionally with a compound rationally designed to be capable of binding a binding pocket of a PSH. . . unknown structure, by generating an X-ray diffraction pattern from a crystallized form of the molecule or molecular complex, using a **molecular replacement** method to interpret the structure of the molecule, where the **molecular replacement** method uses the structural coordinates given in the specification, or its subset comprising a binding pocket, where the structural coordinates. . .

TT TT: NOVEL SYNTHASE HOMOLOGUE **PROTEIN** FUNCTION **PROTEIN** **CRYSTAL** FORM USEFUL IDENTIFY DESIGN INHIBIT ACTIVATE PROTEIN DESIGN ANTIMICROBIAL.

L3 ANSWER 15 OF 22 WPIDS COPYRIGHT 2004 THOMSON DERWENT on STN
AN 2003-221757 [21] WPIDS
CR 2003-229481 [22]
DNC C2003-056524
TI Novel AmB amino transferase protein or its functional **protein** subunit, in a **crystalline** form, useful for identifying and designing inhibitors and activators of the protein, and for designing antimicrobials.
DC B04 D16
IN BADGER, J; BUCHANAN, S G; HENDLE, J; NEWMAN, J; NOLAND, B
PA (BADG-I) BADGER J; (BUCH-I) BUCHANAN S G; (HEND-I) HENDLE J; (NEWM-I) NEWMAN J; (NOLA-I) NOLAND B; (STRU-N) STRUCTURAL GENOMIX INC
CYC 100
PI WO 2003006674 A2 20030123 (200321)* EN 289
RW: AT BE BG CH CY CZ DE DK EA EE ES FI FR GB GH GM GR IE IT KE LS LU MC MW MZ NL OA PT SD SE SK SL SZ TR TZ UG ZM ZW
W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ OM PH PL PT RO RU SD SE SG SI SK SL TJ TM TN TR TT TZ UA UG US UZ VN YU ZA ZM ZW
US 2003105010 A1 20030605 (200339)
AU 2002322445 A1 20030129 (200452)
ADT WO 2003006674 A2 WO 2002-US21937 20020712; US 2003105010 A1 Provisional US 2001-305428P 20010713, US 2002-193858 20020712; AU 2002322445 A1 AU 2002-322445 20020712
FDT AU 2002322445 A1 Based on WO 2003006674
PRAI US 2001-305428P 20010713; US 2002-193858 20020712
TI Novel AmB amino transferase protein or its functional **protein** subunit, in a **crystalline** form, useful for identifying and designing inhibitors and activators of the protein, and for designing antimicrobials.
AB WO2003006674 UPAB: 20040813
NOVELTY - An AmB aminotransferase (AmB) protein (I), or a functional subunit of AmB **protein**, in its **crystalline** form, is new.

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for the following:

(1) producing (M1) a computer. . . binding pocket of a AmB protein, by obtaining 3D structural coordinates defining the protein or a binding pocket of the **protein**, from a **crystal** of the protein, and introducing the structural coordinate into a computer to produce a database containing the molecular structural coordinates. . . produced by M1;

(3) producing (M2) a computer readable database comprising a representation of a binding pocket of a AmB **protein** in a co-**crystal** with a compound, optionally with a compound rationally designed to be capable of binding a binding pocket of a AmB. . . unknown structure, by generating an X-ray diffraction pattern from a crystallized form of the molecule or molecular complex, using a **molecular replacement** method to interpret the structure of the molecule, where the **molecular replacement** method uses the structural coordinates given in the specification, or its subset comprising a binding pocket, where the structural coordinates. .

.

TT TT: NOVEL AMINO TRANSFERASE **PROTEIN** FUNCTION **PROTEIN** **CRYSTAL** FORM USEFUL IDENTIFY DESIGN INHIBIT ACTIVATE PROTEIN DESIGN ANTIMICROBIAL.

L3 ANSWER 16 OF 22 WPIDS COPYRIGHT 2004 THOMSON DERWENT on STN
AN 2003-112007 [10] WPIDS
DNN N2003-089147 DNC C2003-028697
TI Identifying a search model to use in **molecular replacement** for determining a structure of a target biomolecule from crystal data comprises employing computer executable logic.
DC B04 D16 T01
IN ABOLA, E; DAVID, P R; DELFT, F V; MCREE, D; RAMMELKAMP, J; VON DELFT, F
PA (ABOL-I) ABOLA E; (DAVI-I) DAVID P R; (DELF-I) DELFT F V; (MCRE-I) MCREE D; (RAMM-I) RAMMELKAMP J; (SYRR-N) SYRRX INC
CYC 100
PI WO 2002091287 A2 20021114 (200310)* EN 58
RW: AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW MZ NL OA PT SD SE SL SZ TR TZ UG ZM ZW
W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ OM PH PL PT RO RU SD SE SG SI SK SL TJ TM TN TR TT TZ UA UG US UZ VN YU ZA ZM ZW
US 2002183861 A1 20021205 (200310)
AU 2002305345 A1 20021118 (200452)
ADT WO 2002091287 A2 WO 2002-US13988 20020503; US 2002183861 A1 US 2001-848866 20010504; AU 2002305345 A1 AU 2002-305345 20020503
FDT AU 2002305345 A1 Based on WO 2002091287
PRAI US 2001-848866 20010504
TI Identifying a search model to use in **molecular replacement** for determining a structure of a target biomolecule from crystal data comprises employing computer executable logic.
AB WO 200291287 UPAB: 20030211
NOVELTY - Identifying a search model to use in **molecular replacement** for determining a structure of a target biomolecule from crystal data comprises employing computer executable logic.

DETAILED DESCRIPTION - Identifying a search model to use in **molecular replacement** for determining a structure of a target biomolecule from crystal data comprises:

(a) employing computer executable logic to perform multiple **molecular replacement** searches on crystal data of the target biomolecule, where a group of different biomolecule structures are used as search models for the multiple **molecular replacement** searches; and

(b) employing computer executable logic to compare solutions from the multiple **molecular replacement** searches, where the comparison produces data from which biomolecule structures in the group

can be identified as having superior structural. . . medium, useful in association with a computer that includes a processor and a memory, comprising:

(a) logic for performing multiple **molecular replacement** searches on crystal data or diffraction data of a target biomolecule where a group of different biomolecule structures are used as search models for the multiple **molecular replacement** searches; and

(b) logic for comparing solutions from the multiple **molecular replacement** searches.

USE - The method is useful for identifying a search model in **molecular replacement** for determining a structure of a target biomolecule from crystal data (claimed).

Dwg.0/3

TECH UPTX: 20030211

TECHNOLOGY FOCUS - BIOTECHNOLOGY - Preferred Method: Identifying a search model for use in **molecular replacement** for determining a structure of a target biomolecule from crystal data further comprises employing computer executable logic to select the. . . structures used to perform the multiple replacement searches. The biomolecule is a protein, DNA, RNA or a complex comprising a **protein**, DNA or RNA. The **crystal** data is X-ray diffraction data, neutron diffraction crystal data, magnetic crystal data, nuclear magnetic resonance crystal data or mass spectrometry crystal data. **Molecular replacement** is performed using a program comprising AmoRe, BRUTE, COMO (Combined **molecular replacement**) , CNS (Crystallography and NMR System), TNT, GLRF (General locked rotation function program), TRANSF (Translation function program), TF (translation function. . . (Fourier inversion direct to reciprocal space) program) or FFTEXP (Reflection data expanding program), preferably EPMR (a program that finds crystallographic **molecular replacement** solutions using an evolutionary search algorithm), or a **molecular replacement** program comprising an evolutionary algorithm for searching six-dimensional space.

Comparing **molecular replacement** solutions comprises:

(a) comparing figures of merit calculated for the **molecular replacement** solutions;

(b) performing a statistical analysis on figures of merit calculated for the **molecular replacement** solutions;

(c) determining which of the biomolecule structures in the group produced a **molecular replacement** solution whose figure of merit is at least two, three, five or ten standard deviations better than the average figure of merit for **molecular replacement** solutions for the biomolecule structures in the group;

(d) comparing root mean square errors for each **molecular replacement** solution of a probability-weighted average over all possible phase choices;

(e) establishing a background correlation level between the biomolecule structures in the group and the target biomolecule based on the **molecular replacement** solutions and determining which of the biomolecule structures in the group produced a **molecular replacement** solution that exceeds the background correlation level by at least two, three, five or ten standard deviations.

The group of different biomolecule structures on which **molecular replacement** searches are performed comprises:

(a) at least 3 different biomolecule structures, at least one biomolecule structure that has less than. . . comprises a combination of two or more structure fragments. The data produced from the comparison identifies which biomolecule structures produced **molecular replacement** solutions that are at least among the top 35% of **molecular replacement** solutions produced by the group, or that are at least 2, 3, 5 or ten standard deviations better than the **molecular replacement** solutions produced by the group.

Selection of the group of biomolecule structures is:

(a) based, at least in part, on sequence. . . iterative. Selection of the members of the group of biomolecule structures is performed until a biomolecule structure is selected whose **molecular replacement** solution is at least 2, 3, 5 or ten standard deviations better than the **molecular replacement** solution for the biomolecule structures in the group.